# Statistical Simulation

*Postgraduate / MSc Level Lecture Notes*

## 1. Introduction

Statistical simulation refers to the use of random sampling techniques to model and analyze complex systems or stochastic processes. It helps approximate results that are otherwise difficult to compute analytically.

## 2. Motivation

Analytical solutions to real-world problems are often infeasible due to their mathematical complexity. Statistical simulation allows us to understand uncertainty, test hypotheses, and evaluate the performance of statistical estimators.

## 3. Basic Concepts

1 Random number generation using probability distributions

2 Sampling from populations or theoretical distributions

3 Repeated trials to approximate expected values

4 Using simulation outcomes for inference and prediction

## 4. Monte Carlo Simulation

Monte Carlo simulation is based on repeated random sampling to estimate numerical quantities. It is particularly useful when analytical computation is difficult or impossible.

**Example:** Estimation of $\pi$ using random points inside a unit square.

Steps: 1. Generate N random points (x, y) in [0,1] × [0,1]. 2. Count how many fall inside the unit circle ($x^2 + y^2 \leq 1$). 3. Estimate $\pi$ as 4 × (points inside circle) / N.

**Python Example:**

import numpy as np N = 1000000 x = np.random.rand(N) y = np.random.rand(N) pi_est = 4 * np.sum(x**2 + y**2 <= 1) / N print("Estimated π:", pi_est)

## 5. Bootstrapping

Bootstrapping is a resampling technique used to estimate the sampling distribution of a statistic (e.g., mean, median) by drawing samples with replacement from the observed dataset.

**Example:** Estimating confidence interval for mean.

Steps: 1. Draw B bootstrap samples (with replacement) from data. 2. Compute the mean for each sample. 3. Use the distribution of means to estimate standard error and confidence intervals.

**Python Example:**

import numpy as np data = np.array([12, 7, 9, 15, 11, 8, 10]) B = 1000 means = [np.mean(np.random.choice(data, size=len(data), replace=True)) for _ in range(B)] conf_interval = (np.percentile(means, 2.5), np.percentile(means, 97.5)) print("95% CI for mean:", conf_interval)

## 6. Advantages of Statistical Simulation

1. Provides flexibility for complex problems
2. Simplifies uncertainty quantification
3. Enables visual understanding of stochastic processes
4. Applicable across science, finance, and AI

## 7. Limitations

1. Can be computationally intensive
2. Accuracy depends on number of iterations
3. Random number quality affects results

## 8. Summary

Statistical simulation is a powerful tool for approximating complex models using random sampling. Monte Carlo methods and bootstrapping are foundational techniques that provide insights into uncertainty and sampling behavior.

## 9. References


1. Robert, C.P., & Casella, G. (2010). Monte Carlo Statistical Methods.
2. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning.
3. NumPy and SciPy Documentation.